人工智能之于科技期刊出版业态的变革及启示

王亚辉 王 晶
（中国农业科学院农业信息研究所，北京 100081）

摘 要：【目的】探讨人工智能对科技期刊出版业态带来的可能影响。【方法】采用趋势分析和案例论证的方法，对人工智能技术语态下科技期刊传统出版的内容生产、传播生态和价值逻辑进行再剖析和审视。【结果】认为未来的出版业态将更加智能化、高效化、定制化、场景化，具体体现为算法化选题策划、自动化内容创作、智能化文稿处理、个性化发行传播、场景化阅读体验、社群化学术交流；介绍了人工智能应用存在的一些现实瓶颈和困扰，如算法歧视、数据不足和“信息茧房”等，提出了如平衡人机关系、重塑编辑价值的人工智能风险挑战的适应性治理思路。【结论】人工智能对科技出版生态的形塑和再造是全链条、多视角的，需要主动拥抱、积极融入，但也要审慎对待人工智能应用的风险挑战并加强规约治理。

关键词：科技期刊；人工智能；智能化；算法；出版 **中图分类号：**G223 **文献标识码：**A
文章编号：1671-0134（2023）01-052-04 **DOI：**10.19483/j.cnki.11-4653/n.2023.01.008
本文著录格式：王亚辉，王晶 . 人工智能之于科技期刊出版业态的变革及启示 [J]. 中国传媒科技，2023（01）：52-55.

翻阅出版史，从“铅与火”到“光与电”再到“数与网”，每一次重大技术突破和革新都给出版行业带来巨大变化。以人工智能为代表的新一轮技术创新也必将驱动传统科技期刊出版行业发生历史性变革，加快智能出版融合发展进程。

有学者就人工智能与出版行业的融合发展做了一些有益的探究。如王晓光认为，人工智能作为一种颠覆性技术，对内容创作、内容编辑、发行传播等多个出版环节都有潜在影响，而且这种影响难于预估。[1] 武菲菲讨论了人工智能融入内容生产、图书编辑、营销推广和用户需求挖掘等出版环节的可能性，认为人工智能技术可被视为未来出版行业的生态环境。[2] 但多数研究或就广义的出版而言，论述的视角较为宏观，对具体出版媒介的指向性不强；或谈及新闻出版、图书出版及专业教育出版，鲜有涉及科技期刊出版；即便在少有的科技期刊出版相关研究中，大部分是概述性或畅想式讨论人工智能对出版业态或出版流程的潜在影响，对具体案例和实践应用的介绍不足，或是围绕某一流程或某个环节，如选题策划、同行评议、内容审校等展开论证，但研究的内容又略显单一。总体而言，目前对人工智能和科技期刊出版在微观和具象层面特别是涉及全出版流程、全链条环节的融合策略、创新路径、制约瓶颈的研究还比较欠缺。[3-4] 本研究就人工智能之于科技期刊的未来出版逻辑、出版业态全方位、多环节的变革与颠覆进行了全景的呈现、具象的阐述和案例的论证。在透析人工智能或将重塑科技期刊传统出版流程、驱动内容生产、渠道分发和交流分享智能升级的同时，也讨论了科技期刊出版行业应用人工智能仍存在的一些现实瓶颈和困扰，以及适应性治理策略。

1. 科技期刊出版业态的变革

人工智能利用数据挖掘、智能算法、机器学习等技术手段进行数据采集、分析和处理，通过智能感知和智能服务，将在很大程度上改变科技期刊的选题策划、内容创作、文稿处理、发行传播、阅读体验和社群建设等一系列业务流程，塑造出版产业全新的内容生产、加工、分发、服务和消费的生态环境，实现出版业态的智能化、高效化、精准化、定制化、场景化。

1.1 算法化选题策划

在大数据、云计算等技术支持下，人工智能借助数据挖掘算法和深度学习技术抓取、挖掘和分析海量数据，这些数据可能包括目标学科或专业的研究热点和核心话题，以及读者阅读偏好和行为特征（评论、回复、下载、转发等），有助于编辑发现并筛选出更具热点性和前瞻性的内容，从而提升选题策划的高效性、精准性。

目前在出版领域，人工智能参与选题策划取得了一些积极的进展。德国的一家新兴出版商 Inkitt 本是一个电子书写作社区平台，平台上积累了超过 20 万部可供阅读作品的电子书资源，它利用算法模型分析

基金项目：中国农业科学院农业信息研究所 2022 年度科技创新工程项目（项目编号：CAAS-ASTIP-2016-AII），中国农业科学院农业信息研究所公益性科研院所基本科研业务费专项资金（项目编号：JBYW-AII-2021-27）。

读者在线的阅读行为和兴趣偏好，预测电子书的市场潜力，将算法标注为畅销书的电子书选题交由传统出版商完成纸质图书出版。[5] 北大方正电子公司推出出版大数据解决方案，通过对电商、阅读、评论平台和社交网络内容、用户和市场数据去重去噪后的数据管理、可用评价、机器学习、情感分析、用户画像的绘制，帮助编辑发现研究热点，优化出版选题。Springer Nature 开发的 SciGraph 关联开放数据平台，集成了包括期刊、论文、项目、专利、主题、研究人员、科研机构、使用数据在内的高通量、跨领域的内容资源和知识数据，在数据融合、知识发现、内容计算基础上，构建学术知识数据融通关联的大规模知识图谱，基于对知识图谱的智能语义搜索、主题聚合探索分析，有助于掌握学科发展脉络、挖掘特色选题。[6] 美国 Clarivate Analytics 旗下的学术信息检索平台 Web of Science 期刊利用 LDA（Latent Dirichlet Allocation）文档主题生成模型，该模型是一种无监督的机器学习技术，通过对特定领域的学科数据进行分析处理，筛选关键热点，借助算法抽取主题词汇，按显著性自动形成特定主题，为期刊的选题组稿以及热点追踪提供方便。[7]

1.2 自动化内容创作

自动化内容创作较早出现在新闻出版领域，尤其是在国内外主流媒体的新闻报道中，写作机器人得到了大规模的应用，典型如美联社的 Wordsmith 自动撰稿平台、新华社的“快笔小新”、腾讯的 Dream writer。在学术和专业出版领域，2019 年 4 月，Springer Nature 出版了世界上第一本由人工智能写作的学术著作《Lithium-Ion Batteries》，这是由德国法兰克福大学应用计算语言学实验室合作开发的名为“Beta Writer”的算法完成的，它使用了基于相似性的聚类程序，将源文档排序为连贯的章节，并生成简洁的论文摘要，引文则以超链接形式注明，方便读者查阅，通过对锂离子电池领域 53，000 多篇学术文献的跨语料自动摘要，综述了该项技术的前沿进展。[8] 来自美国伦斯勒理工学院等机构的科学家们开发了一款叫作“PaperRobot”的人工智能学术助手，借助自然语言处理和自然语言理解技术对生物医学领域已发表的大量论文深度学习，构建背景知识图，并协助撰写给定主题的论文摘要、关键内容，梳理结论甚至提出进一步的研究建议，对“PaperRobot”生成论文的图灵测试结果表明，机器创作的摘要、结论和研究建议比人工写作有更高的接受度。[9]

值得警醒的是，上述人工智能的尝试并不意味着在不久的将来，机器就能取代人类完成逻辑清晰、论述严谨且有创新发现的研究论文，人工智能现在所做的也仅仅是基于知识数据库或背景知识图对现有的知识、概念进行梳理、链接和聚合，它并没有且短期内也不会创造新的知识和科学发现，而新知识、新发现恰恰是人类不断探索、艰苦攻关、协同创新的智慧结晶。

1.3 智能化文稿处理

在组稿阶段，人工智能能够基于知识图谱、关联分析和文本挖掘技术，从各大文献检索平台和网络数据库中发掘学术资源，筛选相关研究方向的重要专家学者，跟踪他们的研究动态，评估他们的学术影响，及时向他们发出稿约，提升组约稿件的针对性和匹配度。如 Semantic Scholar、Iris.ai、UNSILO、Yewno 等基于人工智能的学术搜索工具，采用机器学习、语义分析等方法，提取文本的含义和关键概念，帮助用户快速理解论文内容，定位、分类、筛选有价值的研究、作者和机构，给选题与组稿带来极大便利。[10–12] 具体如 UNSILO 利用自然语言处理和机器学习技术，分析提炼论文内容，提取作者的主要论点和研究发现，还可抓取 PubMed Central 学术数据库中数百万篇的生物医学论文全文，便于编辑对研究成果有快速、大概的了解。[11] 清华大学计算机科学与技术系的研究团队建立了科技资源大数据分析挖掘与服务平台“AMiner”，该平台囊括了超过 2.3 亿篇学术论文、专利和 1.36 亿位学者，集成了专家档案智能抽取、专家智能搜索、学术大数据融合、学术评价等多项功能，提供了针对科技文献、专家学者和学术活动的强大搜索能力。

在审稿阶段，针对现有的论文查重软件采用逐字匹配查询且不能识别近义词或相似句的问题，一些智能化的学术不端检测工具支持识别整个句子或部分段落，甚至一些反剽窃工具开发了能够发现伪造图像的图表检测功能，帮助更加有效地打击学术抄袭。如 Elsevier 设计的自动化编辑系统 Evise，通过检索和匹配程序与 CrossCheck 数据库的文献进行比对来检查论文的剽窃；来自美国 Syracuse University 的 Daniel Acuna 博士研究团队在 bioRxiv 上发表的一篇文章中介绍了一种机器学习算法，该算法使用基于关键点的检测方法对来自生命科学领域 4 324 本期刊的 76 万篇开放获取论文的 200 多万幅图片进行了检测，发现约有 9% 的图像存在高度重复。[13] 人工智能技术还可智能推荐合适的审稿人。如瑞士 Frontiers Publishers 研发的人工智能评审助手 AIRA，结合内部自定义算法并嵌入 Google、CrossRef 的 iThenticate（文档原创性检查工具）和 Editage 的 Ada（论文自动化评估工具）的功能，快速准确评估稿件质量，还可匹配潜在的同行评审专家，并检查编辑、审稿人和作者之间可能存在的利益冲突。

在编校阶段，大量低端、重复的编辑加工校对工作都可交给人工智能的自动排查和纠错系统来高效完

成，包括稿件字词句段、语法修辞有无错误，名词术语、计量单位的书写是否准确，篇章结构、体例格式是否规范，数理统计方法运用是否恰当。如 Elsevier 的 Aries 审稿系统采用 StatReviewer 软件的 AI 功能核查论文的试验方法、统计数据和研究结论的完整性和可靠性[11]；方正智能辅助审校系统应用机器学习和深度学习技术，拥有分词、实体识别、句法分析、深度语言模型等方法，已初步开发完成易错词、敏感词、不规范名词、连接符、全半角、单位大小写、图表公式序号检查等 11 项功能。

1.4 个性化发行传播

科技期刊在长期的出版实践中汇聚了丰富的研究成果、文献资料等内容资源，以及作者、读者、专家、编委等用户信息，奠定了利用智能算法实现内容精准分发的数据基础。人工智能可以追踪用户在各类数据平台、社交媒体上对学术资源的浏览、下载、评论及分享行为，描绘精准用户画像，基于用户阅读需求和兴趣领域，完成内容的适配性和个性化分发。

TrendMD 是加拿大的一家出版技术服务公司，可提供跨平台相关文章的个性化精准推荐，它通过在合作期刊网站后台安装插件，索引期刊论文的历史元数据，并利用跨平台内容推荐模块的协同过滤技术，根据读者阅读趋向实现合作期刊平台和第三方平台论文的精准推荐，每月通过 TrendMD 平台推荐的文章链接高达 8 亿条，惠及超过 1 亿的读者。超星集团推出“域出版”移动出版平台，为用户提供智慧化的学习资源，其特色的“指纹采集”功能结合用户在平台上的“收藏”“最近浏览”“读书排行”等源数据，可建立个人阅读行为和阅读曲线，并按照阅读曲线推送定制化内容。

1.5 场景化阅读体验

不同于传统出版以图文为主的静态传播，场景时代的媒体传播更加关注内容产品本身为用户创造的沉浸式、可视化体验。VR（虚拟现实）、AR（增强现实）、MR（混合现实）等技术丰富了传统出版的内容呈现样式和用户交互方式，对内容和信息的表达更加深刻、直观，为用户带来了深度沉浸、主动参与的“临场”体验。

VR/AR 技术引入科技期刊出版，也可丰富读者的阅读形式，提升读者的阅读体验，特别是在理工农医类科技论文中仪器设备、实验过程、医科手术的立体再现中有较大的应用优势。《上海大学学报（自然科学版）》探索性地在“三值光学计算机”专题出版中，将期刊论文以动静态 AR 呈现，读者通过智能终端下载纸上 AR 的 App，扫描文中图片，即可在移动端完美展示三值光学计算机实物和体验视频。[14]

1.6 社群化学术交流

利用大数据技术的分类、聚类和关联规则挖掘，借助人工智能的精准推荐和有效匹配，寻找具有共同学术旨趣的业内同行，建立科技期刊的用户社群，通过社群传播打造更加宏大的学术空间，提升科技期刊的平台效益和品牌价值。在学术社群内用户可以实时分享学术动态，交流最新的研究成果，还可以就自己感兴趣的话题获得个性化、精准化的学术资源。

Frontiers 一直坚持社群驱动出版的理念和实践，开发了科研社交平台 Loop，科学家们在平台个人主页上可发布科研成果、学术活动等内容，上传论文、视频、新闻等资源，学者间可相互关注，加入学术小组，向科研同行推荐有价值的学术内容，还可在线组织学术交流。国家新闻出版署出版融合发展（武汉）重点实验室发起的开放科学计划 OSID 平台也有类似的功能，支持期刊社、编辑、作者、读者构建学术交流社区，在学术圈内可就期刊论文、学科热点进行学术讨论、资源共享、开放交流和社交互动。

2. 科技期刊出版应用的瓶颈及困扰

人工智能对科技期刊传统出版的业务重塑和价值创新将催生出版的新业态、新模式、新逻辑。但在科技期刊出版行业完全应用人工智能，仍然存在现实的困难和挑战。具体表现为算法驱动的选题策划、内容生产同质化严重，创造性不足；精准推送可能造成信息窄化；出版数据共享困难，数据壁垒普遍存在。

2.1 算法驱动的选题策划或内容同质化

人工智能依赖机器智慧跟踪、筛选学术研究热点和前沿动态，推动传统出版选题策划方式走向智能高效，但在算法主导下对热点、热词的一味追逐，可能造成相同或相关学科不同刊物选题方向和选题内容同质化，而缺乏差异化、创新性，无法彰显办刊特色，不利于学科创新和出版生态的健康发展。

2.2 人工智能学科专业性、创造性不足

目前人工智能尚处于弱人工智能阶段，还不具备人类意识特有的推理能力和创造能力，也就不具备思考分析并提出观点的能力，自动创作更多适合模板化、程式化的内容生产，由于专业性、创造力的缺失，人工智能还不胜任需要专业研究、深度阐释、复杂求证的科学论文的创作。

2.3 精准推送可能带来信息窄化

基于用户兴趣爱好提供个性化推荐，虽然提升了信息分发效率，但也会带来“信息茧房”问题[15]，从而造成用户接收信息窄化、视野受限、思维固化，这对科技期刊的受众，主要是科研工作者拓宽信息面、知识面，特别是从事跨领域、跨学科学习研究显然是不利的。

2.4 出版数据交互、共享困难

数据是人工智能的“养料”，没有数据支撑，就谈不上人工智能的应用。但在传统出版单位，数据的价值并没有得到足够的重视，对内容生产、内容传播、出版流程以及用户交互的数据积累、挖掘和利用不够，这些数据的潜在价值并没有被发挥出来，而且不同出版单位、文献数据平台之间的数据相互割裂，难以实现开放共享，数据孤岛、数据壁垒广泛存在。

3. 人工智能风险挑战的应对与治理

人工智能在为科技期刊出版业态变革带来一系列新机遇的同时，也不可避免地伴随着一连串新挑战。应对人工智能的风险挑战，不能单从技术层面着手，根本上还是要处理好人工智能与人类的关系问题。

具体到科技期刊出版领域人工智能应用风险挑战的规约与治理，其基本思路是：平衡人机关系，重塑编辑价值，增进协同共治。人工智能新时代下，需要深化对科技期刊出版流程的把控、编辑活动规律的认知、出版伦理和文化价值的坚守，加强前瞻预防，权衡技术利弊，调试价值冲突和伦理困境；公开算法，让算法透明化，增强算法的可解释性、可理解性、可预测性，关注和防范算法偏见、算法权利滥用和“信息茧房”，解决选题功利化、内容模式化、推送单一化等诸多问题；强化编辑的主体意识、责任意识和把关意识，在选题策划、同行评审上发挥编辑在人文关怀和价值判断上的文化引领性和主观能动性，弥补人工智能的技术缺陷和价值失范，构建人工智能技术与编辑伦理和谐共生的出版生态；推进行业出版数据开放共享、标准统一和业务协同；需要加大全民人工智能的教育普及和人才培养力度，尤其是提高编辑人员了解运用人工智能新技术的专业素养。

结语

人工智能对出版生态的变革和再造是全链条、多角度的，是出版生产力的进一步解放和生产关系的深层次变革。需要用发展的眼光、开放的心态主动拥抱人工智能，将人工智能技术积极融入出版产业发展进程中。强化编辑的主导地位和科学引领责任，引导好、利用好人工智能，也要认真评估算法偏见、“信息茧房”、数据越界、侵权追责等人工智能的技术、法律和伦理问题。

参考文献

[1] 王晓光．人工智能与出版的未来 [J]. 科技与出版，2017（11）：4-6.
[2] 武菲菲．人工智能技术与出版行业的融合应用 [J]. 出版广角，2018（1）：26-28.
[3] 代妮，步召德．人工智能对科技期刊出版的影响 [J]. 出版广角，2021（11）：46-48，51.
[4] 张以芳，张宁．全媒体时代人工智能对医学类学术期刊编辑出版业的影响及应对策略 [J]. 中国传媒科技，2020（9）：11-15.
[5] ALBAZAZ A.How Inkitt publishes your books：from preparation to promotion[EB/OL].https：//degfhkjgfjhf.inkitt.com/how-inkitt-publishes-your-books-from-preparation-to-promotion.2017-05-27/2019-12-18.
[6] 宋宁远．面向智慧数据的科学知识图谱构建——以 SciGraph 为例 [J]. 科技与出版，2017（11）：18-19.
[7] 安琪．大数据时代学术期刊发展新探索 [J]. 中国编辑，2017（7）：58.
[8] 环球网．出版机构施普林格·自然开始用机器写书 [EB/OL].https：//smart.huanqiu.com/article/9CaKrnKjvDR.2019-04-03/2019-12-20.
[9] Wang Q，Huang L，Jiang Z，et al. PaperRobot：incremental draft generation of scientific ideas[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence，Italy：Association for Computational Linguistics，2019：1980-1991.
[10] 李媛．人工智能时代的学术期刊数字化传播 [J]. 中国科技期刊研究，2019（11）：1183-1190.
[11] HEAVEN D.AI peer reviewers unleashed to ease publishing grind[J].Nature，2018（7733）：609-610.
[12] 孙赫男．人工智能技术与学术和专业出版的融合发展 [J]. 出版广角，2018（7）：35-36.
[13] ACUNA D E，BROOKES P S，KORDING K P. Bioscience-scale automated detection of figure element reuse[J].bioRxiv，2018.
[14] 赵宇．《上海大学学报（自然科学版）》AR 出版实践 [J]. 编辑学报，2019（1）：74.
[15] 昌沁．新闻传播中人工智能技术造成的伦理失范与对策 [J]. 中国传媒科技，2020（11）：28-30.

作者简介：王亚辉（1986-），男，河南宝丰，博士研究生，助理研究员，中国农业科学院农业信息研究所农业传播与期刊出版事业部编辑部主任，研究方向为期刊编辑出版。

（责任编辑：张晓婧）